DOCUMENT RESUME

ED 228 194 SP 022 031

AUTHOR Yinger, Robert J.; Clark, Christopher M.

TITLE Understanding Teachers' Judgments about Instruction: The Task, the Method, and the Meaning.

INSTITUTION Michigan State Univ., East Lansing. Inst. for Research on Teaching.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
REPORT NO IRT-RS-121
PUB DATE Dec 82
CONTRACT 400-81-0014
NOTE 40p.
AVAILABLE FROM Institute for Research on Teaching, College of Education, Michigan State University, 252 Erickson Hall, East Lansing, MI 48824 ($4.00).
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Cognitive Processes; *Course Evaluation; *Decision Making; Difficulty Level; Elementary School Teachers; Evaluation Criteria; Evaluation Methods; *Evaluative Thinking; Grade 4; Grade 5; Instructional Materials; Intermediate Grades; *Language Arts; *Learning Activities; Skill Development; Student Participation; *Teacher Attitudes; Teaching Methods

ABSTRACT

Six experienced teachers of fourth and fifth grades read and made judgments about the appropriateness, attractiveness, usefulness, and effectiveness of each of 32 short descriptions of language-arts activities. The activity descriptions had been edited to vary systematically on five features: amount of student involvement; difficulty for students; integration of multiple skills or subject matters; demand on the teacher; and fit between stated purpose and instructional process. Each teacher thought aloud during the judgment process; these deliberations were audiotaped and transcribed. A policy-capturing analysis indicated that only small proportions of variance in teacher judgments could be explained in terms of the five manipulated cues. A process-tracing analysis of the protocols of two of the six teachers revealed that they were attending to several activity features in addition to the five mentioned above. The process-tracing analysis also provided a description of how the teachers defined and modified the judgment task and drew upon their own professional experiences to comprehend, edit, and evaluate the activity descriptions. The process-tracing approach appears to be a more promising method of inquiry into complex teacher-judgment tasks than is the policy-capturing approach. (Authors/JD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED228194

Research Series No. 121

UNDERSTANDING TEACHERS' JUDGMENTS ABOUT INSTRUCTION: THE TASK, THE METHOD, AND THE MEANING

Robert J. Yinger

and

Christopher M. Clark

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Christopher M. Clark

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Published by

The Institute for Research on Teaching
252 Erickson Hall
Michigan State University
East Lansing, Michigan 48824-1034

December 1982

This work is sponsored in part by the Institute for Research on Teaching, College of Education, Michigan State University. The Institute for Research on Teaching is funded primarily by the Program for Teaching and Instruction of the National Institute of Education, United States Department of Health, Education, and Welfare. The opinions expressed in this publication do not necessarily reflect the position, policy, or endorsement of the National Institute of Education. (Contract No. 400-81-0014)

SP 022 031

Institute for Research on Teaching

The Institute for Research on Teaching was founded at Michigan State University in 1976 by the National Institute of Education. Following a nationwide competition in 1981, the NIE awarded a second contract to the IRT, extending work through 1984. Funding is also received from other agencies and foundations for individual research projects.

The IRT conducts major research projects aimed at improving classroom teaching, including studies of classroom management strategies, student socialization, the diagnosis and remediation of reading difficulties, and teacher education. IRT researchers are also examining the teaching of specific school subjects such as reading, writing, general mathematics, and science, and are seeking to understand how factors outside the classroom affect teacher decision making.

Researchers from such diverse disciplines as educational psychology, anthropology, sociology, and philosophy cooperate in conducting IRT research. They join forces with public school teachers, who work at the IRT as half-time collaborators in research, helping to design and plan studies, collect data, analyze and interpret results, and disseminate findings.

The IRT publishes research reports, occasional papers, conference proceedings, a newsletter for practitioners, and lists and catalogs of IRT publications. For more information, to receive a list or catalog, and/or to be placed on the IRT mailing list to receive the newsletter, please write to the IRT Editor, Institute for Research on Teaching, 252 Erickson Hall, Michigan State University, East Lansing, Michigan 48824-1034.

Co-Directors: Jere E. Brophy and Andrew C. Porter

Associate Directors: Judith E. Lanier and Richard S. Prawat

Editorial Staff
Editor: Janet Eaton
Assistant Editor: Patricia Nischan

Abstract

Six experienced teachers of fourth and fifth grades read and made judgments about the appropriateness, attractiveness, usefulness, and effectiveness of each of 32 short descriptions of language-arts activities. The activity descriptions had been edited to vary systematically on five features: amount of student involvement, difficulty for students, integration of multiple skills or subject matters, demand on the teacher, and fit between stated purpose and instructional process. Each teacher thought aloud during the judgment process. These deliberations were audiotaped and transcribed. A policy-capturing analysis indicated that only small proportions of variance in teacher judgments could be explained in terms of the five manipulated cues. A process-tracing analysis of the protocols of two of the six teachers revealed that they were attending to several activity features in addition to the five mentioned above. The process-tracing analysis also provided a description of how the teachers defined and modified the judgment task and drew upon their own professional experiences to comprehend, edit, and evaluate the activity descriptions. The authors conclude that the process-tracing approach is a more promising method of inquiry into realistically complex teacher-judgment tasks than is the policy-capturing approach.

UNDERSTANDING TEACHERS' JUDGMENTS ABOUT INSTRUCTION: THE TASK, THE METHOD, AND THE MEANING[1]

Robert J. Yinger and Christopher M. Clark[2]

Sound judgment is a critical skill in any profession, be it law, medicine, engineering, architecture, or teaching. Because of its importance, the judgment process has become the object of psychological research in recent years (Slovic, Fischoff, & Lichtenstein, 1977). In the research on teacher thinking, teacher judgment has emerged as one of several foci of research (Clark & Yinger, 1977; Shavelson & Stearn, 1981). This paper reports the results of a study of teacher judgment during an instructional-activities selection process.

The most frequently used approach to studying and representing judgment processes is policy capturing (Slovic & Lichtenstein, 1971; Shulman & Elstein, 1975). It involves the use of a simple mathematical model (usually linear) and attempts to reproduce the inferential responses of a particular judge. Of central interest in this paradigm is how judges weigh and combine information from cues or features of the objects to be judged.

This approach has been used widely in research on teacher thinking. Recently, policy-capturing models were used to represent teachers'

---

[1]This paper was presented at the annual meeting of the American Educational Research Association in New York City, March 1982.

[2]Robert Yinger is a former IRT researcher with the Teacher Planning Project and an assistant professor of education at the University of Cincinnati. Christopher Clark coordinated the IRT's Teacher Planning Project and is an associate professor of education at Michigan State University.

judgments of the characteristics of effective teachers (Anderson, 1977), effective classroom management (Cone, Note 1), classroom organization (Borko, Note 2), instructional strategies (Russo, Note 3), and instructional content (Floden, Porter, Schmidt, Freeman, & Schwille, 1981).

Process tracing, another approach to modeling problem solving and decision making, has rarely been used to study teacher judgment. Methods involving process-tracing approach very differently the investigation and representation of thinking processes. Since introspective reports of many judges seem to indicate the presence of a complex and configural judgment process, process tracing begins with a complex representation of the judgment using verbal protocols and simplifying the mental processes involved by using decision trees or networks, flow diagrams, and so on.

Process tracing is now commonly used to study problem-solving processes (e.g., Newell & Simon, 1972). It has also proven successful in modeling chess thinking (de Groot, 1965), medical diagnosis (Elstein, Shulman, & Sprafka, 1979), business decision making (Clarkson, 1962), and consumer choice (Bettman, 1971). In research on teaching, Yinger (1980, Note 4) used process-tracing methods to characterize teacher planning.

The study reported here is part of a series of studies investigating teacher judgment during the selection of instructional materials. These studies include one that identifies factors influencing the selection of instructional activities (Clark, Yinger, & Wildfong,

Note 5), a policy-capturing study of teacher judgment (Yinger, Clark, & Mondol, Note 6), a process-tracing study of teacher judgment, a feature analysis of preferred instructional activities, and an analysis of teachers' self-reported judgment processes. This paper presents and discusses a representation of the task of teachers selecting classroom activities (that representation described by process-tracing methods) and compares this representation with that generated by the policy-capturing method.

The major research questions guiding this study are:

1. What factors do teachers take into account when selecting instructional activities?
2. What thinking processes are involved in arriving at these judgments?
3. What are the relative effectivenesses of policy-capturing and process-tracing methods for modeling these types of judgment?

Method

Subject

Six fourth- and fifth-grade teachers from two Michigan school districts volunteered to participate in this study. The three male and three female teachers ranged in age from the mid-twenties to the mid-thirties and averaged 4.8 years of teaching experience (the range was from 4 to 6 years). Four teachers taught in self-contained classrooms; two others team taught. The participants were paid for their participation.

Materials

Thirty-two one-or two-page descriptions of language-arts writing activities were derived from a catalogue of language-arts activities for upper elementary classrooms (Forte, Frank, & McKenzie, 1973). The activity descriptions, presented in the same general format, consisted of an activity title, a one or two sentence statement of the purpose of the activity, and a listing of the steps involved in planning and conducting the activity.

The edited activity description reflected five dimensions found to be important in teachers' judgments of the quality of language-arts instructional materials (Clark, Yinger, & Wildfong, Note 3). These dimensions, or cues, were

1. student involvement,
2. difficulty for students,
3. integration with other skills or subject matter,
4. demand on the teacher, and
5. fit between the stated purpose and described instructional process.

The 32 activity descriptions represented a full factoral matrix of high and low values for each cue. After editing, four researchers examined each activity description to determine whether the cue levels had been set in the intended configuration. Disagreements among the researchers on this task were resolved by further editing and negotiation.

We hypothesized that the evaluation and selection of instructional materials is not a one-judgment process but rather a series of judgments leading to a final decision to implement or not to

implement an activity. (As we will show below, this even proved to be an oversimplified conception of the activities involved in the judgment task.) To investigate this process and to better represent the complexity of the judgment task, each participant was asked to respond to four questions about each activity. On the back of each activity description the questions were stated along with a nine-point continuum to record each response. The four questions or judgments to be made about each activity were as follows:

1. How *attractive* is this activity to you?
2. How *appropriate* is this activity as part of a catalogue of language arts activities for fourth and fifth grade teachers?
3. How likely would you be to *use this activity as it is* in your present classroom?
4. How *effective* do you think this activity would be for your students?

Procedure

In two judgment sessions, each lasting approximately one hour, each teacher first was given information describing the purpose of the study, the procedures, and the materials. Then, a brief background questionnaire and the required consent and payment forms were completed. After a set of six warm-up activities and a question and answer session, each participant was given a loose-leaf booklet containing the 32 activity descriptions (arranged in a different random order for each teacher). The teacher was instructed to proceed through the booklet of activities, completing the first 16 activities in the first judgment session and the remaining activities in the second session.

In addition to responding to the four judgment questions on the reverse side of each activity description, each teacher was instructed to "think aloud" as they participated in the task. These verbalizations were tape recorded and later transcribed, becoming protocols of the judgment task.

Policy-Capturing Analysis

The primary purpose of the policy-capturing analysis was to access the degree to which judgments about instructional activities could be modeled by linear mathematical representations. Simple linear regression equations have been shown to provide, in a number of cases, very good explanations of judgmental responses (e.g., Goldberg, 1968). Based on the past success of this method, linear regression equations were computed for each of the six participants. The five cues were treated as independent variables on which the ratings given to each case were regressed. In all, 24 regression equations were computed using programs from the Statistical Package for the Social Sciences.

Two kinds of information about the judgments emerged. For each participant, regression weights were studied to determine which factors were significant predictors of that judge's rating of the instructional activities. Second, squared multiple correlations were used to estimate the proportion of variance in each judge's ratings explainable in terms of the five cues present in each activity description.

Process-Tracing Analysis

Process-tracing representations are based on records of thinking-in-progress, usually in the form of verbal protocols of the subject's "thinking aloud" during a task. Protocol analysis is a technique that has been devised to infer from these records the information-processing mechanisms underlying human problem solving behavior (Newell, 1977). Akin (1979) has summarized the theoretical conceptions underlying this approach:

> All problem-solving behavior is assumed to consist of transforming a given state of information about the problem into another state such that the second state is closer to containing the information that describes a solution to the problem than the first one. The act of transforming a problem state into another one is called an operation. Given this taxonomy a protocol starts with an initial information state (or the problem description) followed by a sequence of many intermediate states before reaching a final (or solution) state. Each state can be obtained by the application of an operation (or operator) to the previous state. (p. 115)

Within this framework, the activity judgment task can be thought of as a problem-solving task culminating in four judgments related to the attractiveness, appropriateness, usefulness, and effectiveness of each activity. In other words, we are not only interested in the final judgments made by the teachers but also in the thinking process that leads to a judgment.

The traditional procedures used for protocol analysis have been most strongly influenced by the work of Newell and Simon (1972). These researchers have sought to produce formal representations of thought processes that could be simulated by computer models. They

have attempted this, using the problem behavior graph (PBG). A PBG represents a problem-solving task as that of finding the right sequence of operators (or path) to transform the initial state of the problem into the solution state. In its complete form a PBG provides (1) a catalogue of operations that can be applied in a task environment, (2) the circumstances under which such applications are made, and (3) the paths developed during the search for a solution (Akin, 1979; Newell, 1977).

As the use of process-tracing techniques has increased in recent years, a number of criticisms of protocol analysis have been raised (e.g., Nisbett and Wilson, 1977). Four criticisms are asserted most frequently.

1. Studies have shown that when subjects have been asked to introspect about their actions, they are often in error.

2. The verbalization of thought processes interferes with and distorts the normal course of thinking.

3. Since thought processes are much faster than motor responses (e.g., speech), verbalization cannot accurately convey the richness of the cognitive processes.

4. Process-tracing analyses rely upon very small numbers of subjects in each experiment.

Space does not allow an extensive analysis and response to these criticisms here. Briefly, we can say that many of these concerns have been shown to be unfounded. For example, experimental analyses of verbal-reporting processes indicate that verbalization of objects currently in memory can be accurately conveyed without distorting normal thought processes. The innacurate reports cited by some researchers

seem to result from asking subjects to report information that was never directly noticed or used, thus forcing them to infer rather than remember mental processes (see Ericcson and Simon, 1980 for a review of this research).

Missing data (criticism #3) seems to be most problematic when researchers study tasks that require visual and motor processes that are difficult to put into words (e.g., in architectural design). Small sample size in these experiments is greatly offset by the hundreds and often thousands of observations found in the protocols of each subject. Conclusions are generalizations about the nature of the cognitive task and about consistencies within the problem solver rather than generalizations to the population of problem solvers.

Results

Policy-capturing and process-tracing analyses of judgment are most powerful as intra-individual methods. Since both methods generate a sizable amount of individual data that becomes less meaningful when summarized across subjects, we have chosen to present results for only two of six teachers who participated in the study. Data from two subjects will adequately illustrate the types of information generated by the two methods and will be a feasable amount of data to attempt to present and discuss within these space limitations. The two teachers whose judgment processes are described in this report were selected because they have the same number of years of experience

teaching upper-elementary grades and because the proportions of variance in their judgments explained by their policy equations are representative of those of the six participating teachers.

The Policy-Capturing Analysis

Teacher 3, a female, instructs a fourth-grade class in a rural school; she has four years of teaching experience. The mean rating given by this judge to all activity descriptions was 5.26 with a standard deviation of 2.80. The average squared multiple correlation from the four regression models generated for the four judgments was .28 (adjusted R^2 = .14), which indicates that, as a conservative estimate, a little less than one-sixth of the variation in Teacher 3's ratings can be accounted for by the five activity features manipulated in the study.

The portion of this teacher's judgment accounted for by the regression equations shows a fairly consistent weighting pattern across the four judgments. Of the five manipulated cues (Table 1), Teacher 3's judgments of the activity descriptions were influenced most by the perceived amount of student involvement in an activity (Cue 1) followed by demand on the teacher (Cue 4) and the fit of the activity with the stated purpose (Cue 5). This teacher gave the other two cues no significant weighting for any judgment.

Teacher 5, a male, instructs a fifth-grade class in a suburban school; he also has four years of teaching experience. The mean rating given by this judge to all activity descriptions was 6.57 (standard deviation = 2.53). The average squared multiple correlation

Table 1

Regression Weights for Four Judgments
Teachers 3 and 5

Cues	Judgment							
	Attractive		Appropriate		Likely to Use		Effective	
	T3	T5	T3	T5	T3	T5	T3	T5
1. Involvement	.37**	.41	.19*	0	.72*	.19	.40**	.19
2. Difficulty	.19	.16	.06	.25	.34	.50	.28	.19
3. Integration	.06	.22	.06	0	.22	-.37	.16	-.06
4. Demand	.12	-.53	.06	-.13	.72*	-.87	.34*	-.13
5. Fit	.19	-.41	.19*	-.43	.47	-.25	.34*	-.37

*$p < .10$
**$p < .05$

from the four regression models was .08 (adjusted $R^2 = 0$). This indicates that the regression models of this teacher's judgment captured virtually none of the variation in judgment. This is further confirmed in Table 1, which shows that no significant regression equations were generated for this teacher.

The results for these two teachers, though somewhat disturbing, were similar to results from the other four teachers. Of the 24 regression models generated for all six teachers, only 11 equations had significant cue weightings, and a majority of the regression equations had low R^2's (mean $R^2 = .16$; range from .03 to .35). These results are also consistent with another study of 19 teachers judging these same materials (Yinger, Clark, and Mondol, Note 6) in which 44 percent of the regression equations showed no significant regression weights and where the squared multiple correlations were also low (mean $R^2 = .18$; range from .03 to .50).

In the previously cited study, we hypothesized about the reasons for the poor showing of the policy-capturing models for this task. We suspect that a major reason is the limitation of the model in trying to represent a complex naturalistic judgment in terms of only a few cues. The process-tracing analysis that follows will address this hypothesis in more detail.

Process Tracing

The protocol analysis produced three types of information about the judgment activity: (1) a description of the heuristic

operators used by the judges, (2) a reconceptualization of the task itself, and (3) a description of the cues used by the judges during the task. Each of these will be discussed in turn.

Heuristic operators. Processing information during problem-solving tasks has been described in a variety of ways, all of which emphasize the organized, deliberate, and purposeful orientation that people adopt in most cases. Newell and Simon (1972) characterize these processes in terms of operators. Others have used terms such as plans (Miller, Galanter, and Pribram, 1960), schemas (Akin, 1979), and scripts (Abelson, 1976). Often the terms reflect the level of analysis adopted by the researchers. Newell and Simon reduce the process to basic operators that might be simulated on a computer, while those using schemas and scripts are referring to broader organizations of complex, temporal action.

We have chosen to use the term *heuristic operator* to describe the processes used in this study. The word *operator* is used to refer to the manipulations or operations that the judge (teacher) is called upon to do. We preface this by the word *heuristic* to distinguish our use from the very specific operators in the Newell and Simon tradition. Also we see these operators as being employed in the heuristic sense of framing a strategy or approach that is used in a flexible and variable way.

The initial protocol analyses indicated that the teachers were doing distinguishably different things at different points in the task. Some of these activities were repeated regularly, others occurred less frequently. On the basis of this analysis, approximately a dozen

"information processing routines" were initially identified in the two teachers' protocols. By more careful examination of the operators and by subsuming some under others, the list was reduced to seven heuristic operators that were widely used by both teachers. They are reading, interpreting, mental trying-out, categorizing, editing, evaluating, and justifying.

1. Reading. Reading is the process of gathering information from the text of the activity description. Reading was the only source of information available to the judges and thus was a necessary part of the task. The participants read in a variety of ways, including rereading the same material again, scanning ahead to pick up general information (e.g., the length of the description), and "skip reading" (skipping over certain information like lists or examples).

2. Interpreting. Interpreting occurred often and in close proximity to reading. It showed that participants made an effort to understand what was read. Interpreting included paraphrasing previous reading and constructive activity--information mentioned that, while consistent with the text, was not required or suggested by it. A portion of a protocol showing the alternations between reading and interpretation at the beginning of an activity judgment is listed in Figure 1.

3. Mental trying-out. Mental trying-out refers to visualizing an activity in progress while preparing to make a judgment. Basically, it involves mentally playing out or running through an activity

Heuristic Operator	Line	Commentary
Reading	1.	The first activity is number 28 and
	2.	the name of the activity is Hot
	3.	Off the Press. The purpose: After
	4.	completing this activity, the
	5.	student should be able to summarize
	6.	in writing the current events of
	7.	his classroom life and present in
	8.	several styles of writing the
	9.	concepts he is learning.
Interpreting	10.	Ok, so he's going to have to sum-
	11.	marize in writing and he's going
	12.	to be dealing with different
	13.	styles of writing. Ok, so I will
	14.	try and remember those.
	15.	Alright. The first thing he'll
	16.	do is to...
Reading	17.	Introduce the idea of writing a
	18.	newspaper to your class by bring-
	19.	ing lots of newspapers to school
	20.	for examination by the students.
Interpreting	21.	Ok, so this would be one style of
	22.	writing.
Reading	23.	Consider together all the parts
	24.	and purposes of a newspaper, the
	25.	style of writing, methods of
	26.	reporting, etc.
Interpreting	27.	Ok, um..., I don't know if the
	28.	purpose of the activity is to
	29.	worry about the style of writing...
	30.	whether you would need to go into
	31.	the parts and the purposes of the
	32.	newspaper--but that may be
	33.	beneficial--or the methods of
	34.	reporting. I guess it would
	35.	depend on how much you wanted to
	36.	stick to the purpose of the ac-
	37.	tivity.

Figure 1. Reading and interpreting in a protocol.

to gather information about its workability. It often involves picturing the activity as it fits into a wider context.

De Groot (1965) first used this term in his analysis of chess thinking. This term was also used by Yinger (1979, 1980) to characterize how teachers investigate instructional plans. Two distinct types of trying-out were evident in these protocols. The first, projection, involved imagining the activity within the context of the teacher's current modus operandi in the classroom with his/her students. An example from one of the protocols illustrates this operator:

> (Reading) Using three of this week's spelling words, tell a story about the magic potion to the class. Write sentences using these spelling words with correct spelling and punctuation. (Projection) My spelling program is a little different--it's not that the whole class has spelling words. They have their own spelling words. So I guess that would have to be worked out. Good.... We'd come up with varied sentences--which may be very good....

Visualization refers to a more abstract or general trying out where the judge is not "mapping" the activity onto a specific situation but rather trying to visualize the general case. For example:

> I'm trying to think if fifth graders would think that was too juvenile or not....I think (pause)....ok.

Visualization was used less often than projection. It may be that trying-out is harder to do and less fruitful in the abstract.

4. Categorizing. Categorizing, as the name suggests, involves the judges matching the activity description with a familiar activity category or type. The teachers reduced all or portions of

the activity to categories and types of instruction that they were familiar with, relying on pre-stored knowledge and opinions to come to understand judgment. Teachers categorized generally and specifically. They categorized generally when they tried to match the written description with a type of instructional process or activity like role playing, a "practice activity," a "filler activity," or an "application activity." Specific categorization involved a more precise matching of the activity with one the teacher had used before. The following protocol segment exemplifies this operation:

> OK, I've seen this done before--this kind of activity. And I've even done this kind of activity with my own students on different holidays.

5. Editing. Editing is a process of mentally changing in the activity description during the reading and judgment process. Editing most often took place in conjunction with mental trying-out or categorizing. Judges seemed to be strongly disposed toward modifying activities into what they considered more workable forms--a goal that was not a part of the task presented to them (more will be said on this point later).

6. Evaluating. The teachers commonly made evaluative comments throughout the protocols. This seemed to help the teachers prepare for the final judgment ratings by assigning positive or negative weightings to the activity as it was being interpreted and analyzed. In certain places, global, nonspecific evaluations were made, for instance, "I like this!", "This is nice.", or "There's something wrong here." At other times, they made specific or pointedly critical

comments: "I like the idea of proofreaders" or "I don't like the idea of the students being independent."

7. Justifying. Justifying involves presenting a rationale for a decision or evaluative remark. It often accompanied evaluating or editing, most commonly following the responses to the four judgment questions at the end of each activity. Most likely, some justifying was a response to the task constraints (i.e., thinking aloud as part of a study of teacher judgment). But justifying also seemed to help the teacher explain an opinion to him/herself. By seeing how much support could be garnered for a criticism, modification, or evaluation, the judges seemed to be checking their own reasoning processes--a kind of self-testing. A portion of a protocol that illustrates justifying in relation to evaluating and editing is shown in Figure 2.

To begin to understand when and how these heuristic operators were used by the judges, we explored more closely how the teachers approached the activity-rating task. This is the topic of the following section.

Understanding the judgment task.

> Newell and Simon's theory emphasizes that rational human problem solving is characterized by being adapted to the problem to be solved. Since it is adaptive, the behavior of a person solving a problem tells us more about the structure of the task than about the personality dynamics of the problem solver. (Shulman and Elstein, 1975, p. 14)

Process tracing reveals important information about both the task and the thinking processes of those engaged in it. Earlier we said we were interested not only in the judgment process itself

Heuristic Operator	Line	Commentary
Reading	1.	Ok, number three. Let students
	2.	exchange their written work with one
	3.	other student.
Interpreting	4.	Ok, they've written down what
	5.	they're doing.
Reading	6.	Have them read and correct the paper
	7.	for punctuation and spelling.
Interpreting	8.	Ok, I guess we're getting a little
	9.	English in there too. The purpose is
	10.	to identify several creative uses of
	11.	common objects.
Rereading (purpose)	12.	...Several creative uses...and
	13	should gain increased facility in
	14.	the processes required for critical
	15.	thinking and...mature judgments.
Evaluating	16.	Ok, I wouldn't...I don't know if I
	17.	like concentrating so much on the
	18.	written work.
Justifying	19.	Because if you're trying to get
	20.	critical thinking out of it, the
	21.	written work I don't think should
	22.	be emphasized.
Editing	23.	If I used this I wouldn't emphasize
	24.	that.

Figure 2. Evaluating, justifying, and editing in a protocol.

but also in how subjects prepared for judgment. In the next section we discuss our conceptions of the activity judgment task based upon the protocol analysis and related research.

Johnson (1955) characterized judgment as a final component in a thought process preceded by preparation and production processes. These prior considerations become necessary because judgment is based on some type of matching of the objects to be judged with other similiar objects or representations of them (Johnson, 1972). Based on these notions we expected to observe the teachers engaged in some kind of matching behavior as a major part of the task.

We also expected to observe the teachers employing processes to help them understand the written descriptions. Recent problem-solving research has indicated that understanding a problem is an important first step to solving a problem (e.g., Hayes and Simon, 1974; Hayes, Waterman, and Robinson, 1977). Understanding has been defined in this research as the process of constructing representations of a problem or task. (This process is an example of Newell and Simon's notion of constructing a "problem space" within which a solution will be sought.) Representation of a problem based on written information relies upon general knowledge (syntactic and semantic knowledge) but also relies upon knowledge of specific problem types.

Initially, we conceived of the activity-judgment task as a two-step process:

1. understand/represent the activity description; and
2. make a judgment.

It could also be viewed in terms of the subject's goals: judgment (major goal) and representation (sub-goal).

As the analysis progressed we realized that the "understand, then judge" model was too simplistic. First evaluations were made very early for many activities and judgments were being constructed hand-in-hand with the construction of the activity representation. Also, we observed a sizable amount of mental manipulation of the activities as the teachers attempted to construct activity representations and fit them to specific activity routines or types (categorizing) or to compare them against their general conceptions about learning, instruction, and principles of practice (mental trying-out).

Our current conception of the activity-judgment task describes it in terms of the task goals (the goals the teachers set for themselves) and the heuristic operators used to attain the goals. We have divided the task goals into four components:

1. Understand/represent the written description (reading, interpreting, categorizing).
2. Answer the questions, "What would this activity look like in practice (mental trying out)?" "How well would it work (evaluating)?" and "How well do I like it (evaluating)?"
3. Answer the question, "How could I make this activity work (editing, justifying)?"
4. Make final judgments.

This model of the task structure is not intended to be a strict representation of the flow of processing in the task but rather a breakdown of the various operations that seem to occur in the minds

of the teachers. Specifically, Components 2 and 3 help to explain much of the exploration and analysis behavior that was apparent prior to judgment. We think that the four judgment questions (see materials section) created several problems that needed to be solved before an activity could be rated. The most expedient way to solve the problems was through investigating the activities using visualization or projection or through simplifying the new activity to a familiar form (categorization). In fact, there was a distinguishable portion of the protocol for every activity judgment that showed evidence of one or more of these operators.

The third component in the model accounts for the editing behavior that was common to both teachers. The question "How could I make this work?" was not a part of the task, but a self-imposed problem-solving step. In fact, Question 3 (see materials section) for each activity asked the teachers to judge the likelihood that they would use the activity in their classroom "as it is." "How could I make this work?" is possibly an example of the "collector's disposition," the tendency for professionals to collect and store problem representations and solutions without a currently intended use. (See Yinger, Note 4, for a discussion of this behavior in one teacher's planning.) The teachers frequently used editing processes in their deliberations--in 63% of the activities judged by Teacher 5 and in 32% of those of Teacher 3.

The fourth goal, making the final judgments, was a necessary and expected part of the task. Previously we said that we had

observed evaluating and justifying being used earlier in the process. This reinforces our notion that considerable overlap or intertwining of the processes existed in the accomplishment of the task goals, even though it makes sense to talk about them in four parts.

Our representations (Figure 3) of the sequence and nesting of heuristic operators for Teachers 3 and 5 in their analysis of the same activity are analogous to problem behavior graphs (see the section, Policy Tracing Analysis) but are not at a more general level.

Focus of attention in the task. We have thus far described the judgment task as approached by the teachers and the heuristic operators they appeared to use in the process. Now we will describe aspects of the activities or their context that were present in the teachers' thoughts during the activity-judgment task. This analysis is the process-tracing counterpart to the information about cue usage in the policy-capturing analysis and may help to illuminate the results obtained there.

In their judgments of the activity descriptions, both teachers mentioned quite a few considerations or cues. Teacher 3 mentioned 22 different cues and Teacher 5 mentioned 28. Four of the five cues manipulated in the activity descriptions were among those mentioned by each teacher; neither mentioned "integration." When considering a single activity, each used a moderate number of cues (mean cue use per activity: Teacher 3, 6.81; Teacher 5, 4.75). Teachers 3

Heuristic Operators Used By Teacher 5

Reading
  Scanning
    Categorizing
Reading
  Skip-reading
   Evaluating (global)
    Evaluating (specific)
      Editing
       Projecting
        Justifying
  Evaluating (global)
   Categorizing
 Judgment #1
  Justifying
 Judgment #2
  Evaluating (specific)
 Judgment #3
  Justifying
 Judgment #4
  Justifying

Heuristic Operators Used By Teacher 3

Reading
  Visualizing
    Categorizing
      Projecting
Reading
  Evaluating (global)
    Visualizing
  Evaluating (global)
Reading
  Rereading
    Interpreting
      Projecting
    Interpreting
Reading
  Evaluating (global)
Reading
  Evaluating (global)
    Projecting
 Judgment #1
  Justifying
 Judgment #2
  Justifying
 Judgment #3
  Justifying
 Judgment #4

Figure 3. Sequence of heuristic operators employed by two teachers while examining and judging Activity No. 8.

and 5 mentioned factors in their deliberations on the activities that could not always be considered cues of the activity descriptions alone. They considered many aspects of the activities that take on meaning only in the context of implementation. Since mental trying-out places activities in the context of implementation, we observed teachers talking about many specific and often idiosyncratic concerns related to themselves, their students, and their classroom. All of the considerations for both teachers, however, can be grouped under one of three foci: activity, teacher, or student.

Twelve considerations or cues were used by both teachers. In relation to the activity itself, both mentioned fit of the activity process with its stated purpose, clarity of the activity description, activity type, and age level appropriateness of the activity. The cues related to students were (1) student involvement, (2) student enjoyment, (3) difficulty for students, and (4) effective outcomes. Concerns mentioned under teacher focus were (1) demand on the teacher, (2) fit with teacher's goals, (3) fit with current practice ("what I do"), and (4) fit with past practice ("what we've done"). A look at the complete cue lists for the two teachers (Figure 5) shows that each considered many aspects beyond those mentioned in the activity descriptions as a means of accomplishing their task goals.

Process-tracing analysis does not produce a metric comparable to the cue weightings derived from a regression model. We can, however, get a rough indicator of the importance of various considerations

Teacher 3

Focus		
Student	Teacher	Activity
Student involvement	Demand	Fit of purpose and description (3)
Student difficulty	Fit with Teacher's goals (5)	Clarity of procedures
Students' task-related ability (2)	Prerequisite instruction (1)	Appropriateness of instructional strategy
Incidental learning	Fit with current practice	Activity type
Student interest (6)	"Feel"	Internal consistency
Cognitive outcomes	Fit with past practice	Age level appropriateness (4)
Affective outcomes		
Student enjoyment (7)		
General student outcomes		
Individual differences		

Teacher 5

Focus		
Student	Teacher	Activity
Student involvement (5)	Demand	Fit of purpose and description (2)
Student difficulty	Fit with current practice	Brevity
Student enjoyment (1)	Fit with past practice	Variety
Affective outcomes	Fit with teacher's goals (3)	Academic defensibility
Peer interaction	Enthusiasm	Terminology
Student choice		Clarity
Success		Uniqueness (4)
Student needs		Sequence
Challenge		Design/flow
		Diagnostic opportunity
		Activity type
		Practicality
		Expansion potential
		Age-level appopriateness (6)

Figure 5. Cue lists for Teachers 3 and 5.

from the number of times cues were used. (It is important to remember that frequency of use provides no clue as to the importance of a cue in a specific deliberation or how it was used in relation to the other cues considered at the same time.) For Teacher 3, seven cues accounted for approximately 50% of the incidences of cue use (presented in rank order from most frequently mentioned): (1) prerequisite instruction, (2) students' task-related ability, (3) fit of stated purpose with description, (4) age-level appropriateness, (5) fit with teacher's goals, (6) student interest, and (7) student enjoyment. Cues falling into this same category for Teacher 5 (also rank ordered) are (1) student enjoyment, (2) fit of stated purpose with description, (3) fit with teacher's goals, (4) activity uniqueness, (5) student involvement, and (6) age-level appropriateness.

It is worth noting, as we cautioned above, that frequency of use is not directly comparable to cue weighting. For instance, Teacher 5 frequently used the cues of fit and involvement but these cues were never significantly weighted in the regression models for his judgments. Further comparisons of the information generated by the two modeling methods will be taken up in the discussion section.

Discussion

We now discuss what we have learned in this study about (1) teacher judgment of instructional activities, (2) drawing on professional experience, and (3) methods of modeling professional judgment.

Teacher Judgment of Instructional Activities

This study produced information about how six teachers (two of whom are reported on here) approached the activity-judgment task in terms of the problems they confronted, the problem-solving methods they used, and the factors they took into consideration. We modeled the teachers' approaches to the task as a four-component process. It began with attempting to understand the activity description by constructing a mental representation. This representation was then analyzed and judged as the teachers attempted to answer questions about the activity's workability and potential. Efforts were also made to edit and revise the activity into "better" forms.

This approach to the task raises the question of why the teachers engaged in such extensive examination and even editing of the activity descriptions. We propose two answers here. First, it seems most efficient to examine a plan for action under simulated conditions. A related strategy is to see how well an activity description compares to plans that have been successful in the past. Both of these approaches have been observed in other studies of teacher planning (e.g., Yinger, 1980; Clark and Yinger, Note 7).

Second, through investigation and editing, the two teachers reported here seemed to be complicating the process to simplify the product in order to transform the activity into either something they could use or something they could forget. As a result of this, by the end of their participation, they may have picked up a few workable ideas rather than a number of activities that would still require a lot of work to implement.

This study also added to our knowledge about the processes used in teacher judgment. Seven heuristic operators were proposed: reading, interpreting, mental trying-out, categorizing, editing, evaluating, and justifying. We discussed the use of these processing methods as a means of accomplishing the goals in the task.

A third product of the study was a characterization of the types of factors considered during the judgment task. The two teachers mentioned in this paper considered information related to three foci: the activity, the students, and the teacher. Most significantly, the cues mentioned by the teachers in the protocols showed that they were considering a broad range of factors, many of which were not directly related to the activity itself. In other words, the teachers were going beyond the information given and considering idiosyncratic personal and contextual information. This makes sense given the assumed "set toward implementation" that teachers would bring to examining a collection of activities of this sort, but it poses problems for researchers who would attempt to understand and model judgment in the abstract (see methodological discussion below).

Drawing on Professional Experience

Traditionally, most psychological research has been done on naive or inexperienced subjects (e.g., college sophomores). This has also been true of much of the problem-solving research that has used tasks that require little, if any, specialized knowledge or skill. An early exception to this was De Groot's (1965) work on

chess thinking, which has inspired in recent years an interest in expert problem solving and how experience is brought to bear in professional tasks. This study revealed that teachers drew upon their professional knowledge and experience to help them examine and judge activity plans.

The protocol analysis revealed two methods that teachers incorporated into their activity-analyses. In the first, mental trying-out, teachers either visualized portions of or whole activities operating in generalized situations or projected the activity to their current operating context with the students in the classroom. Categorizing was the second operator and involved matching parts or all of the activity description with known types or categories of instructional processes or activities.

Both mental trying-out and categorizing show that these teachers drew heavily on their professional knowledge and experience in this task. This is consistent with results from other studies of expert problem solving in which methods were used similiar to those of this study. Mental trying-out (projection and/or visualization) is similar to the problem-investigation methods observed in chess thinking (De Groot, 1965; Chase and Simon, 1973). The same processes have been observed in teacher planning (Yinger, 1980; Clark and Yinger, Note 7). Simon (1980) has suggested that methods with a strong perceptual component are important skills that professionals develop.

Categorizing, described in this study, is similar to methods found in other studies that draw upon experience as highly organized, large-scale functional units. These experience units--scripts, schemas, production systems, or routines--have been evident in studies of experienced individuals solving physics problems (Simon and Simon, 1978; Larkin, Note 8), solving algebra word problems (Hinsley, Hayes, and Simon, Note 9), understanding stories (Shank and Abelson, 1977), and planning instructional activities (Yinger, 1979). Studies comparing expert problem solving with novice or beginner problem solving have found that novices approach problems piecemeal and are unable to effectively draw upon large scale organizations of knowledge and experience. In terms of this study, we could expect that less experienced teachers categorize, visualize, and project more simply than more experienced teachers do. We would also predict less editing among novices.

Methods of Modeling Professional Judgment

The results of this study support our earlier assertions (Yinger, Clark, and Mondol, Note 6) that policy-capturing methods only inadequately represent complex judgment tasks. These methods have been shown to be much more effective in highly controlled tasks that limit to a small number of unambiguously present cues the amount of information available. We feel that the trade-off that these studies make in reducing ecological validity to gain experimental control is unacceptable for studying professional judgment. The large number of cues revealed in the process tracing illustrates the limitations of trying to understand or model the process using only five cues.

If policy capturing seems like a limited way to determine the important factors considered by a teacher, it says nothing at all about a teacher's conception of the task and the process and methods used to accomplish it. Policy-capturing models assume that judgment proceeds in a fairly straightforward manner by the additive combination of weighted cues. Though these models have proven to be highly predictive in certain judgment situations, we doubt that they have much explanatory value in terms of actual task conceptions or processes.

Process-tracing methods are imperfect. Much of the knowledge that guides our day-to-day or professional activity is tacit and not easily verbalized. Thus, some of the information relevant to describing thought processes remains uncaptured. Still, the naturalistic, even ethnographic orientation of the method seems to be well suited to understanding teacher thinking. We are assuming, of course, that the improvement of professional practice is based on understanding and not merely on prediction and that examining the tasks, the methods, and the meanings created by teachers is an important avenue to arriving at this goal.

Reference Notes

1. Cone, R. Teachers' decisions in managing student behavior: A laboratory simulation of interactive decision-making by teachers. Paper presented at the annual meeting of the American Educational Research Association, Toronto, March 1978.

2. Borko, H. An examination of some factors contributing to teachers' pre-instructional classroom organization and management decisions. Paper presented at the annual meeting of the American Educational Research Association, Toronto, March 1978.

3. Russo, N. Capturing teachers' decision policies: An investigation of strategies for teaching reading and mathematics. Paper presented at the annual meeting of the American Educational Research Association, Toronto, March 1978.

4. Yinger, R.J. A study of teacher planning: Description and theory development using ethnographic and information processing methods. Unpublished doctoral dissertation, College of Education, Michigan State University, 1977.

5. Clark, C.M., Yinger, R.J., & Wildfong, S.C. Identifying cues for use in studies of teacher judgment (Research Series No. 23). East Lansing, Michigan: Institute for Research on Teaching, Michigan State University, 1978.

6. Yinger, R.J., Clark, C.M., & Mondol, M.M. Selecting instructional activities: A policy capturing analysis (Research Series No. 103). East Lansing Michigan: Institute for Research on Teaching, Michigan State University, 1981.

7. Clark, C.M., & Yinger, R.J. Three studies of teacher planning (Research Series No. 55). East Lansing, Michigan: Institute for Research on Teaching, Michigan State University, 1979.

8. Larkin, J.H. Enriching formal knowledge: A model for learning to solve textbook physics problems (CIP Working Paper #409). Pittsburgh: Carnegie-Mellon University, 1980.

9. Hinsley, D.A., Hayes, J.R., & Simon, H.A. From words to equations: Meaning and representation in algebra word problems (CIP Working Paper #331). Pittsburgh: Carnegie-Mellon University, 1976.

References

Abelson, R.P. Script processing in attitude formation and decision making. In J.S. Carroll & J.W. Payne (Eds.), Cognition and social behavior. Hillsdale, New Jersey: Erlbaum, 1976.

Akin, O. Exploration of the design process. Design Methods and Theories, 1979, 13, 115-119.

Anderson, B.C. Differences in teachers' judgment policies for varying numbers of verbal and numerical cues. Organizational Behavior and Human Performance, 1977, 19, 68-88.

Bettman, J.R. The structure of consumer choice processes. Journal of Marketing Research, 1971, 8, 465-471.

Chase, W.G., & Simon, H.A. The mind's eye in chess. In W.G. Chase (Ed.), Visual Information Processing. New York: Academic Press, 1973.

Clark, C.M., & Yinger, R.J. Research on teacher thinking. Curriculum Inquiry, 1977, 7, 279-304.

Clarkson, G.P.E. Portfolio selection: A simulation of trust investment. Englewood Cliffs, New Jersey: Prentice Hall, 1962.

De Groot, A.D. Thought and choice in chess. The Hague: Mouton, 1965.

Elstein, A., Shulman, L.S., & Sprafka, S.A. Medical problem solving: An analysis of clinical reasoning. Cambridge: Harvard University Press, 1979.

Ericsson, K.A., & Simon, H.A. Verbal reports as data. Psychological Review, 1980, 87, 215-251.

Floden, R.E., Porter, A.C., Schmidt, W.H., Freeman, D.J., & Schwille, J.R. Responses to curriculum pressures: A policy-capturing study of teacher decisions about content. Journal of Educational Psychology, 1981, 73, 129-141.

Forte, I., Frank, I., & Mackenzie, J. Kids stuff: Reading and language arts experiences, intermediate-junior high. Nashville, Tennessee: Incentive, 1973.

Goldberg, L.R. Simple models or simple processes?: Some research on clinical judgments. American Psychologist, 1968, 23, 482-498.

Hayes, J.R., & Simon, H.A. Understanding written problem instructions. In L.W. Gregg (Ed.), Knowledge and cognition. Hillsdale, New Jersey: Erlbaum, 1974.

Hayes, J.R., Waterman, D.A., & Robinson, C.S. Identifying relevant aspects of a problem text. Congitive Science, 1977, 1, 297-313.

Johnson, D.M. The psychology of thought and judgment. New York: Harper & Row, 1955.

Johnson, D.M. A systematic introduction to the psychology of thinking. New York: Harper & Row, 1972.

Miller, G.A., Galanter, E., & Pribram, K.H. Plans and the structure of behavior. New York: Holt, Rinehart & Winston, 1960.

Newell, A. On the analysis of human problem solving protocols. In P.N. Johnson-Laird & P.C. Wason (Eds.), Thinking: Readings in cognitive science. New York: Cambridge University Press, 1977.

Newell, A. & Simon, H.A. Human problem solving. Englewood Cliffs, New Jersey: Prentice Hall, 1972.

Nisbett, R.E., & Wilson, T.D. Telling more than we can know: Verbal reports on mental processes. Psychological Review, 1977, 84, 231-259.

Shank, R., & Abelson, R. Scripts, plans, goals, and understanding. Hillsdale, New Jersey: Erlbaum, 1977.

Shavelson, R.J., & Stern, P. Research on teachers' pedagogical thoughts, judgments, decisions, and behavior. Review of Educational Research, 1981, 51, 455-498.

Shulman, L.S., & Elstein, A.S. Studies of problem solving, judgment, and decision making. In F.N. Kerlinger (Ed.), Review of Research in Education (Vol.3), Itaska, Illinois: Peacock, 1975.

Simon, D.P., & Simon, H.A. Individual differences in solving physics problems. In R.S. Siegler (Ed.), Children's thinking: What develops? Hillsdale, New Jersey: Erlbaum, 1978.

Simon, H.A. Problem solving and education. In D.T. Tuma and R. Reif (Eds.), Problem solving and education: Issues in teaching and research. Hillsdale, New Jersey: Erlbaum, 1980.

Slovic, P., Fischoff, B., & Lichtenstein, S. Behavorial decision theory. Annual Review of Psychology, 1977, 28, 1-39.

Slovic, P., & Lichtenstein, S. Comparison of Bayesian and regression approaches to the study of information processing in judgment. Organizational Behavior and Human Performance, 1971, 6, 649-744.

Yinger, R.J. Routines in teacher planning. Theory into Practice, 1979, 18, 163-169.

Yinger, R.J. A study of teacher planning. Elementary School Journal, 1980, 80, 107-127.